REMARKS

Applicants and the undersigned are most grateful for the time and effort accorded the instant application by the Examiner. The Office is respectfully requested to reconsider the rejections presented in the outstanding Office Action in light of the following remarks.

Claims 1-25 were pending in the instant application at the time of the outstanding Office Action. Of these claims, Claims 1, 13 and 25 are independent claims; the remaining claims are dependent claims. Claims 1, 13, and 25 stand rejected under 35 U.S.C. § 102(b) as being anticipated by Richardson et al. (hereinafter "Richardson"). Claims 2, 6-12, 14, and 18-24 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Richardson in view of Kita et al. (hereinafter "Kita"). Claims 3, 4, 15, and 16 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Richardson in view of Kita, and further in view of Miller et al. (hereinafter "Miller"). In light of the following remarks, reconsideration and withdrawal of the present rejections is hereby respectfully requested. It should also be noted, the comments made regarding the present invention in Applicants' previous Amendments remain equally applicable and are, therefore, incorporated by reference as if fully set forth herein.

As indicated in Applicants' disclosure, when test data for a parser is different in nature than the data on which the parser was trained, the performance of a parser will become worse than that of a matched condition. The present invention thus broadly contemplates adapting statistical parsers to new data. In particular, it is assumed that an

initial statistical parser is available and a batch of new data is given. In unsupervised adaptation, however, true parses of the new data are not available. The initial model preferably includes a finite collection of probability mass functions (pmf's). The pmf's are preferably transformed into a new model via Markov matrices. These Markov matrices are preferably obtained by maximizing the likelihood of test data with respect to the decoded parses using the initial model. The adaptation scheme may also be carried out iteratively. (See Page 3, line 15 - Page 4, line 7) The instantly claimed invention thus requires specifically "**adapting the statistical [parsing] model** via employing a mathematical transform". (Claim 1, emphasis added) Similar language appears in the other independent claims. In broad terms, the present invention relates to the adapting of an existing statistical parser into one that fits better new or unseen text data.

As best understood, in contrast to the present invention, Richardson appears to be directed to a method and apparatus for bootstrapping statistical processing into a rule-based natural language processor. This bootstrapping optimizes the operating of a parse that uses lexicon entries to determine possible parts of speech of words and a set of rules to combine words from the input string into syntactic structures. (Abstract) Richardson operates in three modes: a statistics compilation mode, a parsing mode, and a hybrid mode. In the statistics compilation mode, Richardson applies each lexicon entry and rule while parsing a last sample of representative text. Then statistics are compiled based on the success rate of the rules and lexicon entries, either by storing the number of times the rule or lexicon entry produced an entry in a parse tree or by storing a ratio of the number of times an entry was produced to the number of applications of the rule or lexicon entry.

These statistics are normalized (put into the same format) so that they can be compared from rule to rule or entry to entry. The parsing mode applies the rules and lexicon entries until a single syntax tree is formed for the input, thereby not applying all applicable rules and entries as in the statistical compilation mode. The hybrid mode uses the first set of statistics to optimize the operation of the parser while compiling a second set of statistics. (column 4, lines 15-45; column 7, lines 5-25)

In the most recent Office Action the Examiner asserted that adapting the statistical model via employing a mathematical transform was met by "normalizing the statistics for the rules and lexicon entries (Fig. 1, element 103), where it would be necessary for the normalizing step to be carried out by a mathematical transform, and it is further taught that ratio are calculated and stored for the number of times the rule or lexicon entry produces a record in the parse tree". Moreover, it is taught that the facility continues to update the statistics within the parser, (col. 8, lines 6-10), where it would be necessary that any operation performed on a statistical model would be mathematical in nature" (Page 4, lines 5-8)

Applicants respectfully disagree with the Examiner's interpretation of Richardson to the extent that the Examiner indicates, "normalizing the statistics for the rules and lexicon entries, (Fig 1, element 103), where it would be necessary for the normalizing step to be carried out by a mathematical transform" is performed by Richardson and/or teaches or suggests the present invention's "adapting a statistical model via employing a mathematical transform." (Id.) Generally speaking, the mathematical transform used in the present invention is not an arbitrary mathematical operation, but rather it must

maintain the consistency of a probability distribution, i.e., there is a probability distribution before and after transform. There is simply no teaching or suggestion in Richardson that the normalization of the statistics is necessarily carried out by a mathematical transform. Further, normalization is known in the art to be the process of reducing a complex data structure into its simple, stable structures. This is in stark contrast to a mathematical transformation, which is well-known in the art to be an operation, such as a rotation, reflection, or translation in geometry, or operations using linear algebra and explicitly using matrix theory. A transformation involves a change in the spatial or temporal relationships of data. Any alleged transform in Richardson is not an adaptation of the model to better fit new data, it is just a technique used so that the various entries of the model can be compared. As stated in Richardson, "the facility normalizes the compiled statistics, if necessary, so that the statistics for each rule may be compared to the statistic for each other rule and each lexicon entry." (column 4, lines 29-33) Therefore, the normalization step in Richardson has no connection to the present invention. It is thus respectfully submitted that Richardson falls short of the present invention.

Applicants respectfully submit that the applied art does not anticipate the present invention because, at the very least, "[a]nticipation requires the disclosure in a single prior art reference of each element of the claim under construction." W.L. Gore & Associates, Inc. v. Garlock, 721 F.2d 1540, 1554 (Fed. Cir. 1983); see also In re Marshall, 198 U.S.P.Q. 344, 346 (C.C.P.A. 1978).

The Office also rejected certain claims under 35 U.S.C. § 103(a) over Richardson in combination with various references, asserting "it would have been obvious ... to combine the parsing system of Richardson et al. with the Markov calculations as taught by Kita et al." and "it would have been obvious ... [to] combine the parsing system of Richardson et al. with the Markov calculations as taught by Kita et al. and with the probability mass functions of Miller et al.". Applicants respectfully traverse these rejections.

Kita in combination with Richardson does not overcome the deficiencies of Richardson as discussed above. Neither Richardson nor Kita suggest "adapting the statistic parsing model via employing a mathematical transform". (Claim 1, and other independent claims) A 35 U.S.C. § 103(a) rejection requires that the combined cited references provide both the motivation to combine the references and an expectation of success. Further, such a rejection requires that the two combined references are technically combinable. That is, the combination of the two references is technically and practically possible and could be carried out by one of ordinary skill in the art. However, the Richardson and Kita references are not technically combinable. The combination of the two references would not be a valid working invention and thus have absolutely no expectation of success.

More importantly, however, neither Kita nor Richardson address improving an existing statistical parser when applied to newly acquired data by mathematically transforming the existing model. The present invention does not involve Hidden Markov Models (HMM), instead it uses a Markov transform, which is not an equivalent. HMMs

have an underlying Markov chain with probability functions associated with either states or transitions. Since HMMs are not used in the present invention the algorithm estimating or updating HMM parameters in Kita doses not apply to the present invention. Appreciating the fact that HMMs are distinctively different from the present mathematical transform, it is clear the rejections using this reference are without support and should be withdrawn.

Additionally, the Applicant would like to make several other points regarding the present rejections. With regard to claims 2 and 14, as indicated above, the present invention does not use nor claim to use HMMs. Therefore, Kita does not teach or suggest these claims and their rejections should be withdrawn. Per the rejections of claims 6 and 18, as indicated above, Kita, as best understood, uses a well-known technique for estimating HMM parameters, while the present invention, in at least one embodiment, is related to a novel technique for adapting an existing statistical parser to new data. In particular, Kita finds the most optimal probabilities to predict the next phone while parsing. Upon reaching the best probability point, Kita stops the process of parsing.

With regard to Claims 6 and 18, the combination of references does not meet the limitations of the claims in contention. The instant invention, as exemplified in Claims 6 and 18, approaches maximization of the log probabilities in the Markov transform in a much different and distinct fashion. Specifically, the instant invention chooses a Markov matrix such that the log probability of either the decoded parses of test material or adaptation material is maximized. As can be seen, this is separate and distinct from the parsing method of Kita. Further, Kita does not even utilize Markov matrices in this

function; rather, Kita finds these "best" and "highest" probabilities while parsing using a parsing tree (which is known to be separate from a Markov matrix).

With respect to claims 7, 9, 19, and 21, it is respectfully submitted that neither Richardson, nor the combination of Richardson with Kita, teaches both supervised or unsupervised adaptation as claimed in the instant invention. The outstanding Office Action asserts that Kita teaches the use of the Viterbi algorithm to update the probabilities. However, it is well-known in the art that the Viterbi algorithm does not explicitly require supervised and/or unsupervised adaptation. As is well-known in the art, the Viterbi algorithm is used to find the most probable sequence of hidden states given a sequence of observed states for a particular Hidden Markov Model. This has no effect on how the adaptation of data utilizing the HMM is carried out, nor does it require or limit how such adaptation is done. Thus, it is respectfully submitted that using the Viterbi algorithm has no effect on supervised or unsupervised adaptation as taught and claimed in the instant invention. Further, because neither Richardson, Kita, nor the combination of the two, teach supervised or unsupervised adaptation, the rejection with respect to Claims 7, 9, 19, and 21 is respectfully requested to be reconsidered and withdrawn. These claims are respectfully submitted to be allowable over the prior art.

Claims 8 and 20 depend upon Claims 7 and 19. Because Claims 7 and 19 are deemed to be allowable over the prior art, it follows that Claims 8 and 20 would also be allowable subject matter. Thus, the rejection of these claims is also respectfully requested to be reconsidered and withdrawn.

Regarding the rejections of claims 11 and 23, while Richardson addresses decoding test material it fails to teach or suggest decoding for the purpose of adapting an existing parser to new data and/or decoding within a method or process comprising constructing an initial parser and then applying the proposed adaptation technique. As to the rejection of claims 10 and 22, "an efficient parsing mode, where the parser only applies applicable rule and lexicon entries" fails to relate to adapting an existing model to new material. These rejections should now be withdrawn.

Miller et al. in combination with Richardson and Kita does not overcome the deficiencies of Richardson or Kita as discussed above. Further, the combination of these references is also not technically valid, as shown above with respect to Richardson and Kita. A 35 U.S.C. § 103(a) rejection requires that the combined cited references provide both the motivation to combine the references and an expectation of success. Therefore, it is respectfully submitted that the rejections based upon the combination of Miller with Richardson and Kita are not valid rejections. Reconsideration and withdrawal is respectfully requested. The Applicants would like to also note that none of these references, including Miller, address the process of transforming an original probability mass function into another, since it must be understood that there are two (2) probability mass functions, one occurring before and one occurring after adaptation. Furthermore, whether the probability mass function is written as a row or column vector is secondary to the novelty of the use of a mathematical transform as presently included in the claims.

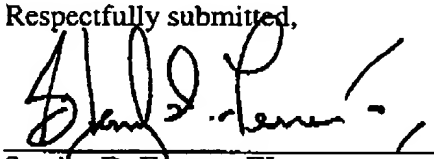
By virtue of dependence from what are believed to be allowable independent Claims 1 and 13, it is respectfully submitted that Claims 2-12 and 14-24 are also

presently allowable. Applicants acknowledge that Claims 5 and 17 were indicated by the Examiner as being allowable if rewritten in independent form. Applicants reserve the right to file new claims of such scope at a later date that would still, at that point, presumably be allowable.

The "prior art made of record" has been reviewed. Applicants acknowledge that such prior art was not deemed by the Office to be sufficiently relevant as to have been applied against the claims of the instant application. To the extent that the Office may apply such prior art against the claims in the future, Applicants will be fully prepared to respond thereto.

In summary, it is respectfully submitted that the instant application, including Claims 1-25, is presently in condition for allowance. Notice to the effect is earnestly solicited. If there are any further issues in this application, the Examiner is invited to contact the undersigned at the telephone number listed below.

Respectfully submitted,



Stanley D. Ference III
Registration No. 33,879

Customer No. 35195
FERENCE & ASSOCIATES
409 Broad Street
Pittsburgh, Pennsylvania 15143
(412) 741-8400
(412) 741-9292 - Facsimile

Attorneys for Applicants